EXPRESS MAIL NUMBER: EL 904 925 641 US DATE OF DEPOSIT: August 13, 2001

I hereby certify that this paper is being deposited with the United States Postal Service "EXPRESS MAIL Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to: Box PATENT APPLICATION, Commissioner for Patents; Washington, DC 20231.

Mary R. Zimmerman

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application for:

Hieu Van TRAN et al.

Serial No.: Not yet assigned

Filing Date: August 13, 2001

For: **ARRAY ARCHITECTURE AND OPERATING METHODS FOR DIGITAL MULTILEVEL NONVOLATILE MEMORY INTEGRATED CIRCUIT SYSTEM**

Examiner: Not yet assigned

Group Art Unit: Not yet assigned

PRELIMINARY AMENDMENT

Box Patent Application
Commissioner for Patents
Washington, DC. 20231

Sir:

Prior to examination of the above-identified application, which is a divisional of application serial number 09/231,928 filed on January 14, 1999, Applicants request the following amendments and cancellation and amendment of claims be entered.

IN THE SPECIFICATION:

Please add the following heading and paragraph at page 1, after the title and before the first heading:

--CROSS REFERENCE TO RELATED APPLICATIONS

This application is a divisional of and claims the benefit of U.S. Application No. 09/231,928, filed January 14, 1999, the disclosure of which is incorporated herein by reference.--

Please replace the paragraph beginning at page 1, line 15, with the following rewritten paragraph:

--As the information technology progresses, the demand for high density giga bit and tera bit memory integrated circuits is insatiable in emerging applications such as data storage for photo quality digital film in multi-mega pixel digital camera, CD quality audio storage in audio silicon recorder, portable data storage for instrumentation and portable personal computers, and voice, data, and video storage for wireless and wired phones and other personal communicating assistants.--

Please replace the paragraph beginning at page 7, line 26, with the following rewritten paragraph:

--A cell structure of one typical SSI flash cell is symbolically shown in Fig. 1A. Its corresponding transistor symbol is shown in Fig. 1B. The cell is made of two polysilicon gates (abbreviated as poly), a floating gate poly FG 100F and a control gate poly CG 100C. The control gate CG 100C also acts as a select gate that individually select each memory cell. This has the advantage of avoiding the over erase problem which is typical of stacked gate CHE flash cell. The floating gate has a poly tip structure that points to the CG 100C, this is to enhance the electric field from the FG 100F to the CG 100C which allows a much lower voltage in FN erase without using a thin interpoly oxide. The thicker interpoly oxide leads to a higher reliability memory cell. The cell is also fabricated such that a major portion of the FG 100F overlaps the source junction 100S. This is to make a very high

coupling ratio from the source 100S to FG 100F, which allows a lower erase voltage and is advantageous to the SSI programming which will be described shortly. A structural gap between the FG 100F and CG 100C is also advantageous for the efficient SSI programming.--

Please replace the paragraph beginning at page 9, line 34, with the following rewritten paragraph:

--The challenges associated with putting together a billion transistors on a single chip without sacrificing performance or cost are tremendous. The challenges associated with designing consistent and reliable multilevel performance for a billion transistors on a single chip without sacrificing performance or cost are significantly more difficult. The approach taken here is based on the modularization concept. Basically everything begins with a manageable optimized basic unitary block. Putting appropriate optimized unitary blocks together makes the next bigger optimized block..--

Please replace the paragraph beginning at page 30, line 25, with the following rewritten paragraph:

--High data rate, meaning high sense speed and write speed, is required for data intensive application. The speed is proportional to capacitance and voltage swing and inversely proportional to the current,

$$T = C * V / I \quad (2).--$$

Please replace the paragraph beginning at page 33, line 35, with the following rewritten paragraph:

--In an embodiment as shown in Fig. 5A, lines 240, 241, and 242 are in the middle, sandwiched between lines 240A, 240B, 241A and 241B in the bottom and CL0 264 in the top. Furthermore, line 240 is on top of the spacing between lines 240A and 240B and line 241 is on top of the spacing between lines 241A and 241B. This has the benefit of reducing significantly the bottom plane capacitance of line 240 and line 241 since the oxide below each line is almost

doubled. The lines 240 and 241 could be positioned on top of lines 240A and 241A respectively when the sidewall capacitance reduction outweighs the benefit of the bottom plane capacitance reduction. The sidewall capacitance refers to the capacitance resulting from the vertical walls of a line, the bottom plane capacitance refers to the capacitance from the bottom of a line, and the top plane capacitance refers to the capacitance from the top of a line.--

Please replace the paragraph beginning at page 35, line 29, with the following rewritten paragraph:

--Two approaches are shown in Fig. 5C to reduce this transient phenomenon. In one embodiment, 2-step ramp rate control approach greatly reduces this transient effect without prolonging the programming time as follows. First VCL ramps fast during TRP1 to an intermediate voltage VCLINT, e.g., 2-6 V, then VCL stays at an intermediate voltage for a finite time TVCLINT, then VCL ramps slow during TRP2 to a final voltage VCLFIN. The first fast ramp with the flat intermediate time TVCLINT will let transient current flowing through the cell to stabilize most of the cell capacitances such as CBL in a short time and at sufficiently low VCL voltage so that insignificant programming takes place while the transient current is flowing. The TRP1 is made fast to consume little programming time. The second slow ramp then brings the cell to a final programming voltage without affecting the programming rate since very little current is flowing through the cell in the second ramp.--

Please replace the paragraph beginning at page 38, line 20, with the following rewritten paragraph:

--Another factor that is reduced greatly is the excessive leakage current from the bitline to ground due to junction leakage, bitline to bitline leakage, band-to-band tunneling, and cell subthreshold conduction. For example, for a typical leakage to 10 pA per cell, with 16,384 cells per bitline, the total leakage is 164 nA, which is greater than Ipcell = 100 nA. This implies that the multilevel programming will be controlled due to the uncontrollable excessive leakage current contributing to the controlled programming current Ipcell. With the inhibit and

segmentation scheme, the total leakage current is reduced to 128 X 10 pA = 1.28 nA, which is much less than Ipcell = 100 nA.--

Please replace the paragraph beginning at page 39, line 3, with the following rewritten paragraph:

--Fig. 4D shows an alternative array architecture in which a set of inhibit select line INHBLA1-3 and INHBLB1-3 275 to 280 are used to inhibit all segment bitlines except the selected segment bitline. VINH 999 is shared for all the segments. The operating method makes use of a segment cascoding scheme that is described as follows. To even isolate the bitline capacitance further, bitline select transistors 220-227 are also used as cascoding transistors in programming in addition to the select and inhibit function. In programming, cell 200 for example, the voltage on line 261 is initially pulsed high to pass inhibit voltage VINH 999 from a page select PSELS 120S into the selected segment bitline SBL0 240A. Then the voltage on line ENBLA0 261 is pulsed to a cascoding voltage VPBCAS, e.g., 1 V. A precharge signal then charges the selected top bitline BLP0 240 to 0.3V. The final voltage on the top bitline BLP0 240 is $=\sim$ 0.3 V since 1 V – VT $=\sim$ 0.3 V. Hence the voltage on line BLP0 240 no longer changes during programming. The voltage on the segment bitline, however, still changes as VCL is applied and stabilized. But the capacitance on the segment bitline is minimal, $=\sim$ 0.15pF. Here the operating method just described could also apply to the array shown in Fig. 4A but the inhibit voltages on the unselected segment bitlines are floating. The array shown in Fig. 4D just makes sure all the unselected segment bitlines are kept at a constant inhibit voltage VINH 999.--

Please replace the paragraph beginning at page 57, line 9, with the following rewritten paragraph:

--Fig. 19A shows various voltages generated and used in one embodiment of the invention for program verifying, program upper and lower margin verifying, read sensing and restore high or restore low margin verifying during read sensing. The read sensing is advantageously performed in the voltage-mode but other modes of read sensing are also applicable. All the voltages are generated by the V&IREF block 172. VREFR(L) is the program

verify voltage used to verify program level L of a reference cell. VREFD(L) is the program verify voltage used to verify program level L of a data cell. For example, in a 4 bit per cell storage embodiment there are 16 levels used. It is also possible to use 15 levels instead of 16 levels since the extreme low or high levels not need to be constrained to exact low or high levels but can go to ground or power supply respectively. VREFR0 through VREFR15 are program verify voltages used for verifying programming of the reference cells. VREFD0 through VREFD15 are program verify voltages used for verifying programming of the data cells. VUM(L) and VLM(L) are upper and lower program margin voltages respectively for level L. Each level L may have its own VUM(L) and VLM(L) voltage values. VUM(L) and VLM(L) can each be of different value also for each level L. On the other hand, VUM(L) and VLM(L) can be of the same voltage value for all the levels. VUM(L) and VLM(L) voltages are generated by the block V&IREF 172. VRSTH(L) and VRSTL(L) are RESTORE HIGH and RESTORE LOW margin voltages respectively for level L. Each level L may have its own VRSTH(L) and VRSTL(L) voltage value. VRSTH(L) and VRSTL(L) can each be of different value also for each level L. On the other hand, VRSTH(L) and VRSTL(L) can be of the same voltage value for all the levels. VRSTH(L) and VRSTL(L) voltages are generated by the V&IREF 172 block. VCELLR(L) is the voltage read back from a reference cell during read sensing. VCELLD(L) is the voltage read back from a data cell during read sensing. The cross-hatched regions show the distribution of possible read back voltages during read sensing after reference cells or data cells have been programmed to a certain level L, while using VREFR(L) or VREFD(L) as the program verify voltage, respectively. The distributions occur because every cell does not have the same programming or read sensing characteristics.--

Please replace the paragraph beginning at page 59, line 27, with the following rewritten paragraph:

--Fig. 21 shows the flow diagram after page programming begins. The Program flag = Pass is set and the BUSY signal is set. The program inhibit mode of all cells in the page being programmed are reset to enable programming. Based on the output B[0:3] of the data latches of each YDRVS 110S, SYDRVS 114S or RYDRVS 112S a program verify voltage VREFD(L) is set at the input of the comparator in each of the respective YDRVS 110S, SYDRVS 114S or

RYDRVS 112S. Based on the output B[0:3] of the data latches of each REFYDRVS 116S a program verify voltage VREFR(L) is set at the input of the comparator in each REFYDRVS 116S. For each reference cell and data cell in the page being programmed, the cell voltage VCELLD(L) or VCELLR(L) is read. Depending on the output B[0:3] of the data latches (a) for each REFYDRVS 116S the appropriate program verify voltage VREFR(L) is compared to the reference cell read back voltage VCELLR(L) and (b) for each YDRVS 110S, SYDRVS 114S, RYDRVS 112S, the appropriate program verify voltage VREFD(L) is compared with data cell read back voltage VCELLD(L) to indicate whether further programming is required. If no further programming is required for a particular reference cell or data cell, it is put in the program inhibit mode. If the Program Pulse Count = MAXPC is not true, then the cells are placed in the program mode and another programming pulse is applied to all the cells in the page, including the reference cells. Cells which are in the program inhibit mode do not get any additional programming. Cells which are not in the program inhibit mode get additional programming. After the programming pulse is applied, the program pulse count is incremented and the cells are placed in the voltage-mode read to verify if further programming is required. This iterative verify-program loop is continued until either all the cells in the page including the reference cells are in the program inhibit mode or when the program pulse count = MAXPC is true. If program pulse count = MAXPC true condition is reached, before all cells in the page including the reference cells are all in the program inhibit mode, then the program flag = fail condition is set, BUSY signal is reset and the programming cycle is done. Whenever the All Cells in Program Inhibit Mode = true condition is reached, the flow moves to the next step as shown in Fig. 22A.--

Please replace the paragraph beginning at page 61, line 25, with the following rewritten paragraph:

--Fig. 23 shows the flow diagram for the page read cycle. During a page read cycle a plurality of memory cells are read in parallel. However this algorithm is equally applicable for single cell read. After the page read command is issued along with the address of the page to be read, the BUSY signal is set, RESTOREL and RESTOREH flags are reset, the data latches in the YDRVS 110S, SYDRVS 114S, RYDRVS 112S are set to output B [0:3] = 1111 and N is set to

3. N represents the number of bits stored per memory cell. All the cells in the addressed page are placed in the voltage-mode read and the cell voltages, VCELLR(L) for reference cells and VCELLD(L) for data cells are read. BN is forced to "0" and the read verify voltage VCELLR(L), which is one of the reference read back voltages dependent on B3, B2, B1, B0, is compared with the cell read back voltage VCELLD(L). For each cell, if the VCELLD(L) > VCELLR(L) then BN is latched as "1", otherwise BN is latched as "0". The loop continues until al the bits B3, B2, B1, B0 are latched and N = 0. Next, as shown in Fig. 24, for each level L, a MARGIN RESTORE LOW Voltage VRSTRL(L) = VCELLR(L) – VRSTL(L) is generated, where VRSTL(L) is the restore low margin voltage. Depending on the latched bits B3, B2, B1, B0 on each of the YDRVS 110S, SYDRVS 114S, RYDRVS 112S, the voltage VRSTRL(L) is compared with the respective data cell read back voltage VCELLD(L). If VCELLD(L) > VRSTRL(L) for any one of the cells, then the RESTOREL flag is set. Next, for each level L a MARGIN RESTORE HIGH Voltage VRSTRH(L) = VCELLR(L-1) + VRSTH(L) is generated, where VRSTH(L) is the restore high margin voltage. Depending on the latched bits B3, B2, B1, B0 on each of the YDRVS 110S, SYDRVS 114S, RYDRVS 112S, the voltage VRSTRH(L) is compared with the respective data cell read back voltage VCELLD(L). If VCELLD(L) < VRSTRH(L) for any one of the cells, then the RESTOREH flag is set, otherwise RESTOREH flag is not set. Next, as shown in Fig. 25, BUSY signal is reset and the byte count ND is initialized to NDI. NDI is the byte count of the existing byte address location. All bits in the respective YDRVSs, SYDRVSs, or RYDRVSs data latches are ready to be sequentially read. Whenever READ CLOCK = Y, the RED_ADD_TRUE is checked for that byte address location. If RED_ADD_TRUE = Y, then data from RYDRVS 112S is output to the IO port IO [0:7] 1001, otherwise data from YDRVS 110S is output to the IO port IO[0:7] 1001. If READ CLOCK = N and ENABLE = Y then the flow loops back until READ CLOCK = Y or ENABLE = N. After all the data is output i.e. ND > MAXND = Y or if ENABLE = N, the Page read cycle is done. If ND > MAXND is = N, then ND is incremented and the flow loops back to check the READ CLOCK.--

Please replace the paragraph beginning at page 67, line 13, with the following rewritten paragraph:

--If at the end of any verify-program iteration, the COMPOR 32 line goes high, the ALGOCNTRL 164 sequences to the margin verify mode. All latches 59 are reset. All cells are placed in the voltage read mode by READB 52 at logic low. At this time inhibit voltage is de-coupled from BLIN 35 and current bias transistor N4 50 is coupled to BLIN 35. Cell voltages VCELLR(L) or VCELLD(L) are respectively available on BLIN 35 of a REFYDRVS 116S or BLIN 35 of YDRVS 110S, SYDRVS 114S or RYDRVS 112S. During program margin verify the voltages read back from the data cells are checked for adequate margin from voltages read back from reference cells for each programmed level L. In the Upper Program Margin Verify mode, voltages UMV(0) through UMV(15) are placed on the VR0 through VR(15). Depending on the output B3, B2, B1, B0 of the data latches DATALAT3 10, DATALAT2 11, DTALAT1 12, DATALAT0 13, within each YDRVS 110S, SYDRVS 114S, RYDRVS 112S one specific voltage UMV(0) through UMV(15) is output to the input VROUT 55 of the VOLTAGE COMPARATOR 27. At this time the VOLTAGE COMPARATOR 27 compares the voltages at its inputs. If voltage on BLIN 35 is higher than voltage on VROUT 55 the output COMPOUT 58 is low, otherwise it is high. At this time a positive going strobe pulse is applied to the ENLATCOMP 29 common to all the latches 59 in YDRVS 110S, SYDRVS 114S and RYDRVS 112S, to latch the status of line COMPOUT 58. If COMPOUT 58 is low, then the COMLATQ 40 remains at logic low. If COMPOUT 58 is high, then the COMLATQ 40 switches to logic high. At this time, if ALGOCNTRL 164 latches a logic low in the status latch in INPUT LOGIC 160 block by looking at the status of the COMPORB 33 line, then a program fail condition is reached and the AlGOCNTRL 164 goes out of the page programming cycle. Otherwise, ALGOCNTRL 164 sequences to the Lower Program Margin Verify mode.--

IN THE CLAIMS:

Please cancel claims 1-3 without prejudice.

Please add the following new claims 4-53.

--4. A data storage system comprising:

a plurality of memory arrays, each memory array including a plurality of memory cells, a plurality of bit lines, a plurality of control gate lines, and at least one common line, wherein each memory cell is configurable to store one of 2^N values where N is two or greater;

at least one memory decoder coupled to the memory arrays, each memory decoder configured to provide bias signals to selected ones of the plurality of memory cells; and

a reference array operatively coupled to the memory arrays and configurable to provide reference signals used for programming and reading the selected ones of the plurality of memory cells.

5. The system of claim 4, wherein the reference array comprises a plurality of reference cells, and wherein the reference cells and memory cells selected for programming are biased with approximately similar bias conditions on their control gates, common lines, and bit lines.

6. The system of claim 5, wherein the each set of reference cell and associated memory cell shares the same common line and control gate line.

7. The system of claim 4, further comprising:

a first address decoder operatively coupled to the plurality of bit lines, the first address decoder configured to receive an input address and select one or more bit lines.

8. The system of claim 4, further comprising:

a second address decoder operatively coupled to the at least one common line, the second address decoder configured to receive an input address and select one or more common lines.

9. The system of claim 4, further comprising:

a third address decoder operatively coupled to the plurality of control gate lines, the third address decoder configured to receive an input address and select one or more control gate lines.

10. The system of claim 4, further comprising:

a reference generator coupled to the at least one memory decoder and the reference array, the reference generator configured to provide a first set of signals to the reference array and the bias signals to the at least one memory decoder.

11. The system of claim 10, furthering comprising:

a voltage multiplier configured to receive an input voltage and generate an output voltage, wherein the output voltage is higher than the input voltage and is used for a combination of erase, program, and read operations.

12. The system of claim 11, wherein the voltage multiplier is implemented using switched capacitor circuits.

13. The system of claim 11, furthering comprising:

a voltage control circuit coupled to voltage multiplier and reference generator, the voltage control circuit configured to provide a reference signal from the reference generator or a high voltage to the memory cells.

14. The system of claim 13, wherein the voltage control circuit is configured to provide the high voltage to a control gate for an erase operation.

15. The system of claim 13, wherein the voltage control circuit is configured to provide a high voltage pulse to a common line for a programming operation.

16. The system of claim 4, wherein each memory array includes

a plurality of segments, each segment including P rows by Q columns of memory cells.

17. The system of claim 16, wherein the plurality of bit lines comprise

a set of main bit lines that traverses a length of each memory array, and

sets of segmented bit lines, each segmented bit line traversing a portion of the length of the memory array and coupled to one main bit line at one or more selected locations.

18. The system of claim 17, further comprising:

a plurality of bit line select transistors coupling the segmented bit lines to the main bit lines.

19. The system of claim 18, wherein the bit line select transistors are also used as cascoding transistors to isolate bit line capacitance.

20. The system of claim 16, wherein the at least one common line comprises

at least one main common line that traverses the width of the memory array, and

a plurality of segmented common lines for each main common line, each segmented common line traversing a portion of the width of the memory array and coupled to one main common line at one or more selected locations.

21. The system of claim 4, further comprising:

a plurality of current sinks disposed along each of the at least one common line.

22. The system of claim 16, further comprising:

a set of main control gate lines that traverses a width of the memory array, each main control gate line coupled to a set of memory cells located along a row of the memory array.

23. The system of claim 22, further comprising:

a set of segmented control gate lines for each control gate line, each segmented control gate line traversing a portion of the width of the memory array.

24. The system of claim 16, further comprising:

at least one inhibit line coupled to memory cells within the memory arrays.

25. The system of claim 24, wherein each inhibit line couples to memory cells in two or more segments.

26. The system of claim 24, wherein each inhibit line couples to memory cells within one memory array.

27. The system of claim 16, further comprising:

a plurality of inhibit enable lines, one inhibit enable line for each bit line.

28. The system of claim 4, wherein each common line is driven from both sides of a memory array.

29. The system of claim 4, wherein the reference array includes

a plurality of reference cells, each reference cell operative to provide one reference signal.

30. The system of claim 29, wherein the reference cells are disposed at approximately linearly spaced locations along the at least one common line.

31. The system of claim 29, wherein the reference cells are disposed at approximately geometrically spaced locations along the at least one common line.

32. The system of claim 29, further comprising:

a plurality of reference lines coupled to the plurality of reference cells, wherein resistance of the reference lines is approximately matched to resistance of common lines.

33. The system of claim 32, wherein each reference cell is disposed at a location in an associated reference line such that voltage drops at both ends of the reference line are approximately equal.

34. The system of claim 29, wherein each reference signal is generated by averaging outputs from two or more reference cells.

35. The system of claim 4, wherein the reference array includes:

a plurality of reference cells operative to provide the reference signals, wherein at least one of the reference signals is generated by extrapolating outputs from two reference cells.

36. The system of claim 4, wherein the reference signals define 2^N unique levels used for programming and reading the memory cells.

37. The system of claim 4, wherein the reference signals define 2^N-1 unique levels used for programming and reading the memory cells.

38. The system of claim 5, wherein the reference cells are programmed with a set of reference values.

39. The system of claim 38, wherein reference cells previously programmed are inhibited from further programming.

40. The system of claim 4, further comprising:

a plurality of drivers coupled to the plurality of bit lines, each driver configured to control an associated bit line during write, read, or erase operation.

41. The system of claim 40, wherein each driver comprises:

a plurality of N data latches configured to receive and latch N data bits from a memory cell during a read operation.

42. The system of claim 40, wherein each driver further comprises:

a voltage comparator coupled to the associated bit line and a reference source, the voltage comparator configured to compare a voltage on the bit line and one of the reference signals from the reference array and to provide a comparison result.

43. The system of claim 42, wherein each driver further comprises:

a multiplexer operative to receive the reference signals from the reference array and to provide one of the reference signals to the voltage comparator.

44. The system of claim 40, wherein each driver is associated with one memory cell during programming, and wherein each driver further comprises:

an inhibit circuit operative to enable or inhibit programming of a particular memory cell coupled to the associated bit line.

45. The system of claim 40, wherein each driver is associated with one memory cell during programming, and wherein each driver further comprises:

control circuitry configured to generate a first status signal indicative of a particular memory cell coupled to the associated bit line being placed in a program inhibit mode.

46. The system of claim 45, wherein the plurality of drivers are configured to generate a second status signal indicative of all memory cells associated with the plurality of drivers being placed in the program inhibit mode.

47. The system of claim 46, wherein the second status signal is generated by wired-ORing first status signals from the plurality of drivers.

48. The system of claim 46, wherein the first status signals from the plurality of drivers are each provided to a gate of a respective one of a plurality of pull-down transistors, and wherein the second status signal is generated by coupling drains of the pulldown transistors.

49. The system of claim 40, wherein each driver includes control circuitry configured to generate a third status signal indicative of a particular memory cell coupled to the associated bit line being incompletely programmed.

50. The system of claim 49, wherein the plurality of drivers are configured to generate a fourth status signal indicative of at least one memory cells associated with the plurality of drivers being incompletely programmed.

51. The system of claim 50, the fourth status signal is generated by wired-ORing third status signals from the plurality of drivers.

52. The system of claim 50, wherein the third status signals from the plurality of drivers are each provided to a gate of a respective one of a plurality of pull-down transistors, and wherein the fourth status signal is generated by coupling drains of the pulldown transistors.

53. The system of claim 4, wherein voltages for memory cells not selected for programming are set to approximately zero.--

IN THE DRAWINGS:

Please amend Figs. 4A-4C and 14 as indicated by the redlined revisions. Applicants submit that no new matter has been added.

REMARKS

This is a divisional application of application serial number 09/231,928 filed on January 14, 1999. Claims 1 through 3 have been cancelled. Claims 4 through 53 have been added.

Consideration of the present application, as preliminarily amended, is respectfully requested.

Please address all future communications regarding this application to:

Edward B. Weller
GRAY CARY WARE & FREIDENRICH LLP
1755 Embarcadero Road
Palo Alto, CA 94303-3340

Please direct all telephone calls to Edward B. Weller at (650) 833-2436

Attached hereto is a marked up version of the changes made to the specification and claims by the current amendment. The attached page is captioned "Version with Markings to Show Changes Made".

Please charge any additional fees, including any fees necessary for extensions of time, or credit overpayment to Deposit Account No. **07-1896, referencing 2102397-991720**.

Dated: August 13, 2001

GRAY CARY WARE & FREIDENRICH LLP

By _____
Edward B. Weller
Reg. No. 37,468

GRAY CARY WARE & FREIDENRICH LLP
1755 Embarcadero Road
Palo Alto, CA 94303-3340
Telephone: (650) 833-2000
Facsimile: (650) 320-7401

Version with Markings to Show Changes Made

In the Specification:

Following the title but prior to the first heading on page 1, insert the following heading and paragraph:

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a divisional of and claims the benefit of U.S. Application No. 09/231,928, filed January 14, 1999, the disclosure of which is incorporated herein by reference.

Paragraph beginning at line 15 of page 1 has been amended as follows:

As the information technology progresses, the demand for high density giga bit and tera bit memory integrated circuits is insatiable in emerging applications such as data storage for photo quality digital film in multi-mega pixel digital camera, CD quality audio storage in audio silicon recorder, portable data storage for instrumentation and portable personal computers, and voice, data, and video storage for wireless and wired phones and other personal communicating assistants.

Paragraph beginning at line 26 of page 7 has been amended as follows:

A cell structure of one typical SSI flash cell is symbolically shown in Fig. 1A. Its corresponding transistor symbol is shown in Fig. 1B. The cell is made of two polysilicon gates (abbreviated as poly), a floating gate poly FG 100F and a control gate poly CG 100C. The control gate CG 100C also acts as a select gate that individually select each memory cell. This has the advantage of avoiding the over erase problem which is typical of stacked gate CHE flash cell. The floating gate has a poly tip structure that points to the CG 100C, this is to enhance the electric field from the FG 100F to the CG 100C which allows a much lower voltage in FN erase without using a thin interpoly oxide. The thicker interpoly oxide leads to a higher reliability memory cell. The cell is also fabricated such that a major portion of the FG 100F overlaps the source junction 100S. This is to make a very high

coupling ratio from the source 100S to FG 100F, which allows a lower erase voltage and is advantageous to the SSI programming which will be described shortly. A structural gap between the FG 100F and CG 100C [CG] is also advantageous for the efficient SSI programming.

Paragraph beginning at line 34 of page 9 has been amended as follows:

The challenges associated with putting together a billion transistors on a single chip without sacrificing performance or cost are tremendous. The challenges associated with designing consistent and reliable multilevel performance for a billion transistors on a single chip without sacrificing performance or cost are [is] significantly more difficult. The approach taken here is based on the modularization concept. Basically everything begins with a manageable optimized basic unitary block. Putting appropriate optimized unitary blocks together makes the next bigger optimized block.

Paragraph beginning at line 25 of page 30 has been amended as follows:

High data rate [rata], meaning high sense speed and write speed, is required for data intensive application. The speed is proportional to capacitance and voltage swing and inversely proportional to the current,

$$T = C * V / I \quad (2).$$

Paragraph beginning at line 35 of page 33 has been amended as follows:

In an embodiment as shown in Fig. 5A, lines [line] 240, 241, and 242 are in the middle, sandwiched between lines 240A, 240B, 241A and 241B in the bottom and CL0 264 in the top. Furthermore, line 240 is on top of the spacing between lines 240A and 240B and line 241 is on top of the spacing between lines 241A and 241B. This has the benefit of reducing significantly the bottom plane capacitance of line 240 and line 241 since the oxide below each line is almost doubled. The lines 240 and 241 could be positioned on top of lines 240A and 241A respectively when the sidewall capacitance reduction outweighs the benefit of the bottom plane capacitance reduction. The sidewall capacitance refers to the capacitance resulting from the vertical walls of a line, the bottom plane capacitance refers to the capacitance from the bottom of a line, and the top plane capacitance refers to the capacitance from the top of a line.

Paragraph beginning at line 29 of page 35 has been amended as follows:

Two approaches are shown in Fig. 5C to reduce this transient phenomenon. In one embodiment, 2-step ramp rate control approach greatly reduces this transient effect without prolonging the programming time as follows. First VCL ramps fast during TRP1 to an intermediate voltage VCLINT, e.g., 2-6 V, then VCL stays at an intermediate voltage for a finite time TVCLINT, then VCL ramps slow during TRP2 to a final voltage VCLFIN. The first fast ramp with the flat intermediate time TVCLINT will let transient current flowing through the cell to stabilize most of the cell capacitances such as CBL in a short time and at sufficiently low VCL voltage so that insignificant programming takes place while the transient current is flowing. The TRP1 is made fast to consume little programming time. The second slow ramp then brings the cell to a final programming voltage without affecting [effecting] the programming rate since very little current is flowing through the cell in the second ramp.

Paragraph beginning at line 20 of page 38 has been amended as follows:

Another factor that is reduced greatly is the excessive leakage current from the bitline to ground due to junction leakage, bitline to bitline leakage, band-to-band tunneling, and cell subthreshold conduction. For example, for a typical leakage to 10 pA per cell, with 16,384 cells per bitline, the total leakage is 164 nA, which is greater than Ipcell = 100 nA. This implies that [Meaning] the multilevel programming will be controlled due to the uncontrollable excessive leakage current contributing to the controlled programming current Ipcell. With the inhibit and segmentation scheme, the total leakage current is reduced to 128 X 10 pA = 1.28 nA, which is much less than Ipcell = 100 nA.

Paragraph beginning at line 3 of page 39 has been amended as follows:

Fig. 4D shows an alternative array architecture in which a set of inhibit select line INHBLA1-3 and INHBLB1-3 275 to 280 are used to inhibit all segment bitlines except the selected segment bitline. VINH 999 is shared for all the segments. The operating method makes use of a segment cascoding scheme that is described as follows. To even isolate the bitline capacitance further, bitline select transistors 220-227 are also used as cascoding transistors in programming in addition to the select and inhibit function. In programming, cell 200 for

example, the voltage on line 261 is initially pulsed high to pass inhibit voltage VINH 999 from a page select PSELS 120S into the selected segment bitline SBL0 240A. Then the voltage on line ENBLA0 261 is pulsed to a cascoding voltage VPBCAS, e.g., 1 V. A precharge signal then charges the selected top bitline BLP0 240 to 0.3V. The final voltage on the top bitline BLP0 240 is =~ 0.3 V since 1 V – VT =~ 0.3 V. Hence the voltage on line BLP0 240 no longer changes during programming. The voltage on the segment bitline, however, still changes as VCL is applied and stabilized. But the capacitance on the segment bitline is minimal, =~ 0.15pF. Here the operating method just described could also apply to the array shown in Fig. 4A but the inhibit voltages on the unselected segment bitlines are floating. The array shown in Fig. 4D just makes sure all the unselected segment [segments] bitlines are kept at a constant inhibit voltage VINH 999.

Paragraph beginning at line 9 of page 57 has been amended as follows:

Fig. 19A shows various voltages generated and used in one embodiment of the invention for program verifying, program upper and lower margin verifying, read sensing and restore high or restore low margin verifying during read sensing. The read sensing is advantageously performed in the voltage-mode but other modes of read sensing are also applicable. All the voltages are generated by the V&IREF block 172. VREFR(L) is the program verify voltage used to verify program level L of a reference cell. VREFD(L) is the program verify voltage used to verify program level L of a data cell. For example, in a 4 bit per cell storage embodiment there are 16 levels used. It is also possible to use 15 levels instead of 16 levels since the extreme low or high levels not need to be constrained to exact low or high levels but can go to ground or power supply respectively. VREFR0 through VREFR15 are program verify voltages used for verifying programming of the reference cells. VREFD0 through VREFD15 are program verify voltages used for verifying programming of the data cells. VUM(L) and VLM(L) are upper and lower program margin voltages respectively for level L. Each level L may have its own VUM(L) and VLM(L) voltage values [value]. VUM(L) and VLM(L) can each be of different value also for each level L. On the other hand, VUM(L) and VLM(L) can be of the same voltage value for all the levels. VUM(L) and VLM(L) voltages are generated by the block V&IREF 172. VRSTH(L) and VRSTL(L) are RESTORE HIGH and RESTORE LOW margin voltages respectively for level L. Each level L may have its own VRSTH(L) and VRSTL(L) voltage

value. VRSTH(L) and VRSTL(L) can each be of different value also for each level L. On the other hand, VRSTH(L) and VRSTL(L) can be of the same voltage value for all the levels. VRSTH(L) and VRSTL(L) voltages are generated by the V&IREF 172 block. VCELLR(L) is the voltage read back from a reference cell during read sensing. VCELLD(L) is the voltage read back from a data cell during read sensing. The cross-hatched regions show the distribution of possible read back voltages during read sensing after reference cells or data cells have been programmed to a certain level L, while using VREFR(L) or VREFD(L) as the program verify voltage, respectively. The distributions occur because every cell does not have the same programming or read sensing characteristics.

Paragraph beginning at line 27 of page 59 has been amended as follows:

Fig. 21 shows the flow diagram after page programming begins. The Program flag = Pass is set and the BUSY signal is set. The program inhibit mode of all cells in the page being programmed are reset to enable programming. Based on the output B[0:3] of the data latches of each YDRVS 110S, SYDRVS 114S or RYDRVS 112S a program verify voltage VREFD(L) is set at the input of the comparator in each of the respective YDRVS 110S, SYDRVS 114S or RYDRVS 112S. Based on the output B[0:3] of the data latches of each REFYDRVS 116S a program verify voltage VREFR(L) is set at the input of the comparator in each REFYDRVS 116S. For each reference cell and data cell in the page being programmed, the cell voltage VCELLD(L) or VCELLR(L) is read. Depending on the output B[0:3] of the data latches [of] (a) for each REFYDRVS 116S the appropriate program verify voltage VREFR(L) is compared to the reference cell read back voltage VCELLR(L) and (b) for each YDRVS 110S, SYDRVS 114S, RYDRVS 112S, the appropriate program verify voltage VREFD(L) is compared with data cell read back voltage VCELLD(L) to indicate whether further programming is required. If no further programming is required for a particular reference cell or data cell, it is put in the program inhibit mode. If the Program Pulse Count = MAXPC is not true, then the cells are placed in the program mode and another programming pulse is applied to all the cells in the page, including the reference cells. Cells which are in the program inhibit mode do not get any additional programming. Cells which are not in the program inhibit mode get additional programming. After the programming pulse is applied, the program pulse count is incremented and the cells are placed in the voltage-mode read to verify if further programming is required.

This iterative verify-program loop is continued until either all the cells in the page including the reference cells are in the program inhibit mode or when the program pulse count = MAXPC is true. If program pulse count = MAXPC true condition is reached, before all cells in the page including the reference cells are all in the program inhibit mode, then the program flag = fail condition is set, BUSY signal is reset and the programming cycle is done. Whenever [When ever] the All Cells in Program Inhibit Mode = true condition is reached, the flow moves to the next step as shown in Fig. 22A.

Paragraph beginning at line 25 of page 61 has been amended as follows:

Fig. 23 shows the flow diagram for the page read cycle. During a page read cycle a plurality of memory cells are read in parallel. However this algorithm is equally applicable for single cell read. After the page read command is issued along with the address of the page to be read, the BUSY signal is set, RESTOREL and RESTOREH flags are reset, the data latches in the YDRVS 110S, SYDRVS 114S, RYDRVS 112S are set to output B [0:3] = 1111 and N is set to 3. N represents the number of bits stored per memory cell. All the cells in the addressed page are placed in the voltage-mode read and the cell voltages, VCELLR(L) for reference cells and VCELLD(L) for data cells are read. BN is forced to "0" and the read verify voltage VCELLR(L), which is one of the reference read back voltages dependent on B3, B2, B1, B0, is compared with the cell read back voltage VCELLD(L). For each cell, if the VCELLD(L) > VCELLR(L) then BN is latched as "1", otherwise BN is latched as "0". The loop continues until al the bits B3, B2, B1, B0 are latched and N = 0. Next, as shown in Fig. 24, for each level L, a MARGIN RESTORE LOW Voltage VRSTRL(L) = VCELLR(L) – VRSTL(L) is generated, where VRSTL(L) is the restore low margin voltage. Depending on the latched bits B3, B2, B1, B0 on each of the YDRVS 110S, SYDRVS 114S, RYDRVS 112S, the voltage VRSTRL(L) is compared with the respective data cell read back voltage VCELLD(L). If VCELLD(L) > VRSTRL(L) for any one of the cells, then the RESTOREL flag is set. Next, for each level L a MARGIN RESTORE HIGH Voltage VRSTRH(L) = VCELLR(L-1) + VRSTH(L) is generated[.], where VRSTH(L) is the restore high margin voltage. Depending on the latched bits B3, B2, B1, B0 on each of the YDRVS 110S, SYDRVS 114S, RYDRVS 112S, the voltage VRSTRH(L) is compared with the respective data cell read back voltage VCELLD(L). If VCELLD(L) < VRSTRH(L) for any one of the cells, then the RESTOREH flag is set, otherwise

RESTOREH flag is not set. Next, as shown in Fig. 25, BUSY signal is reset and the byte count ND is initialized to NDI. NDI is the byte count of the existing byte address location. All bits in the respective YDRVSs, SYDRVSs, or RYDRVSs data latches are ready to be sequentially read. Whenever READ CLOCK = Y, the RED_ADD_TRUE is checked for that byte address location. If RED_ADD_TRUE = Y, then data from RYDRVS 112S is output to the IO port IO [0:7] 1001, otherwise data from YDRVS 110S is output to the IO [io] port IO[0:7] 1001. If READ CLOCK = N and ENABLE = Y then the flow loops back until READ CLOCK = Y or ENABLE = N. After all the data is output i.e. ND > MAXND = Y or if ENABLE = N, the Page read cycle is done. If ND > MAXND is = N, then ND is incremented and the flow loops back to check the READ CLOCK.

Paragraph beginning at line 13 of page 67 has been amended as follows:

If at the end of any verify-program iteration, the COMPOR 32 line goes high, the ALGOCNTRL 164 sequences to the margin verify mode. All latches 59 are reset. All cells are placed in the voltage read mode by READB 52 at logic low. At this time inhibit voltage is de-coupled from BLIN 35 and current bias transistor N4 50 is coupled to BLIN 35. Cell voltages VCELLR(L) or VCELLD(L) are respectively available on BLIN 35 of a REFYDRVS 116S or BLIN 35 of YDRVS 110S, SYDRVS 114S or RYDRVS 112S. During program margin verify the voltages read back from the data cells are checked for adequate margin from voltages read back from reference cells for each programmed level L. In the Upper Program Margin Verify mode, voltages UMV(0) through UMV(15) are placed on the VR0 through VR(15). Depending on the output B3, B2, B1, B0 of the data latches DATALAT3 10, DATALAT2 11, DTALAT1 12, DATALAT0 13, within each YDRVS 110S, SYDRVS 114S, RYDRVS 112S one specific voltage UMV(0) through UMV(15) is output to the input VROUT 55 of the VOLTAGE COMPARATOR 27. At this time the VOLTAGE COMPARATOR 27 compares the voltages at its inputs. If voltage on BLIN 35 [55] is higher than [then] voltage on VROUT 55 the output COMPOUT 58 is low, otherwise it is high. At this time a positive going strobe pulse is applied to the ENLATCOMP 29 common to all the latches 59 in YDRVS 110S, SYDRVS 114S and RYDRVS 112S, to latch the status of line COMPOUT 58. If COMPOUT 58 is low, then the COMLATQ 40 remains at logic low. If COMPOUT 58 is high, then the COMLATQ 40 switches to logic high. At this time, if ALGOCNTRL 164 latches a logic low in the status latch

in INPUT LOGIC 160 block by looking at the status of the COMPORB 33 line, then a program fail condition is reached and the AlGOCNTRL 164 goes out of the page programming cycle. Otherwise, ALGOCNTRL 164 sequences to the Lower Program Margin Verify mode.

In the Claims:

Claims 1 through 3 have been cancelled.

Claims 4 through 53 have been added.

4. (New) A data storage system comprising:

a plurality of memory arrays, each memory array including a plurality of memory cells, a plurality of bit lines, a plurality of control gate lines, and at least one common line, wherein each memory cell is configurable to store one of 2^N values where N is two or greater;

at least one memory decoder coupled to the memory arrays, each memory decoder configured to provide bias signals to selected ones of the plurality of memory cells; and

a reference array operatively coupled to the memory arrays and configurable to provide reference signals used for programming and reading the selected ones of the plurality of memory cells.

5. (New) The system of claim 4, wherein the reference array comprises a plurality of reference cells, and wherein the reference cells and memory cells selected for programming are biased with approximately similar bias conditions on their control gates, common lines, and bit lines.

6. (New) The system of claim 5, wherein the each set of reference cell and associated memory cell shares the same common line and control gate line.

7. (New) The system of claim 4, further comprising:

a first address decoder operatively coupled to the plurality of bit lines, the first address decoder configured to receive an input address and select one or more bit lines.

8. (New) The system of claim 4, further comprising:

a second address decoder operatively coupled to the at least one common line, the second address decoder configured to receive an input address and select one or more common lines.

9. (New) The system of claim 4, further comprising:

a third address decoder operatively coupled to the plurality of control gate lines, the third address decoder configured to receive an input address and select one or more control gate lines.

10. (New) The system of claim 4, further comprising:

a reference generator coupled to the at least one memory decoder and the reference array, the reference generator configured to provide a first set of signals to the reference array and the bias signals to the at least one memory decoder.

11. (New) The system of claim 10, furthering comprising:

a voltage multiplier configured to receive an input voltage and generate an output voltage, wherein the output voltage is higher than the input voltage and is used for a combination of erase, program, and read operations.

12. (New) The system of claim 11, wherein the voltage multiplier is implemented using switched capacitor circuits.

13. (New) The system of claim 11, furthering comprising:

a voltage control circuit coupled to voltage multiplier and reference generator, the voltage control circuit configured to provide a reference signal from the reference generator or a high voltage to the memory cells.

14. (New) The system of claim 13, wherein the voltage control circuit is configured to provide the high voltage to a control gate for an erase operation.

15. (New) The system of claim 13, wherein the voltage control circuit is configured to provide a high voltage pulse to a common line for a programming operation.

16. (New) The system of claim 4, wherein each memory array includes a plurality of segments, each segment including P rows by Q columns of memory cells.

17. (New) The system of claim 16, wherein the plurality of bit lines comprise a set of main bit lines that traverses a length of each memory array, and sets of segmented bit lines, each segmented bit line traversing a portion of the length of the memory array and coupled to one main bit line at one or more selected locations.

18. (New) The system of claim 17, further comprising:
a plurality of bit line select transistors coupling the segmented bit lines to the main bit lines.

19. (New) The system of claim 18, wherein the bit line select transistors are also used as cascoding transistors to isolate bit line capacitance.

20. (New) The system of claim 16, wherein the at least one common line comprises
at least one main common line that traverses the width of the memory array, and
a plurality of segmented common lines for each main common line, each segmented common line traversing a portion of the width of the memory array and coupled to one main common line at one or more selected locations.

21. (New) The system of claim 4, further comprising:
a plurality of current sinks disposed along each of the at least one common line.

22. (New) The system of claim 16, further comprising:
a set of main control gate lines that traverses a width of the memory array, each main control gate line coupled to a set of memory cells located along a row of the memory array.

23. (New) The system of claim 22, further comprising:

a set of segmented control gate lines for each control gate line, each segmented control gate line traversing a portion of the width of the memory array.

24. (New) The system of claim 16, further comprising:

at least one inhibit line coupled to memory cells within the memory arrays.

25. (New) The system of claim 24, wherein each inhibit line couples to memory cells in two or more segments.

26. (New) The system of claim 24, wherein each inhibit line couples to memory cells within one memory array.

27. (New) The system of claim 16, further comprising:

a plurality of inhibit enable lines, one inhibit enable line for each bit line.

28. (New) The system of claim 4, wherein each common line is driven from both sides of a memory array.

29. (New) The system of claim 4, wherein the reference array includes

a plurality of reference cells, each reference cell operative to provide one reference signal.

30. (New) The system of claim 29, wherein the reference cells are disposed at approximately linearly spaced locations along the at least one common line.

31. (New) The system of claim 29, wherein the reference cells are disposed at approximately geometrically spaced locations along the at least one common line.

32. (New) The system of claim 29, further comprising:

a plurality of reference lines coupled to the plurality of reference cells, wherein resistance of the reference lines is approximately matched to resistance of common lines.

33. (New) The system of claim 32, wherein each reference cell is disposed at a location in an associated reference line such that voltage drops at both ends of the reference line are approximately equal.

34. (New) The system of claim 29, wherein each reference signal is generated by averaging outputs from two or more reference cells.

35. (New) The system of claim 4, wherein the reference array includes:

a plurality of reference cells operative to provide the reference signals, wherein at least one of the reference signals is generated by extrapolating outputs from two reference cells.

36. (New) The system of claim 4, wherein the reference signals define 2^N unique levels used for programming and reading the memory cells.

37. (New) The system of claim 4, wherein the reference signals define 2^N-1 unique levels used for programming and reading the memory cells.

38. (New) The system of claim 5, wherein the reference cells are programmed with a set of reference values.

39. (New) The system of claim 38, wherein reference cells previously programmed are inhibited from further programming.

40. (New) The system of claim 4, further comprising:

a plurality of drivers coupled to the plurality of bit lines, each driver configured to control an associated bit line during write, read, or erase operation.

41. (New) The system of claim 40, wherein each driver comprises:

a plurality of N data latches configured to receive and latch N data bits from a memory cell during a read operation.

42. (New) The system of claim 40, wherein each driver further comprises:

a voltage comparator coupled to the associated bit line and a reference source, the voltage comparator configured to compare a voltage on the bit line and one of the reference signals from the reference array and to provide a comparison result.

43. (New) The system of claim 42, wherein each driver further comprises:

a multiplexer operative to receive the reference signals from the reference array and to provide one of the reference signals to the voltage comparator.

44. (New) The system of claim 40, wherein each driver is associated with one memory cell during programming, and wherein each driver further comprises:

an inhibit circuit operative to enable or inhibit programming of a particular memory cell coupled to the associated bit line.

45. (New) The system of claim 40, wherein each driver is associated with one memory cell during programming, and wherein each driver further comprises:

control circuitry configured to generate a first status signal indicative of a particular memory cell coupled to the associated bit line being placed in a program inhibit mode.

46. (New) The system of claim 45, wherein the plurality of drivers are configured to generate a second status signal indicative of all memory cells associated with the plurality of drivers being placed in the program inhibit mode.

47. (New) The system of claim 46, wherein the second status signal is generated by wired-ORing first status signals from the plurality of drivers.

48. (New) The system of claim 46, wherein the first status signals from the plurality of drivers are each provided to a gate of a respective one of a plurality of pull-down transistors, and wherein the second status signal is generated by coupling drains of the pulldown transistors.

49. (New) The system of claim 40, wherein each driver includes control circuitry configured to generate a third status signal indicative of a particular memory cell coupled to the associated bit line being incompletely programmed.

50. (New) The system of claim 49, wherein the plurality of drivers are configured to generate a fourth status signal indicative of at least one memory cells associated with the plurality of drivers being incompletely programmed.

51. (New) The system of claim 50, the fourth status signal is generated by wired-ORing third status signals from the plurality of drivers.

52. (New) The system of claim 50, wherein the third status signals from the plurality of drivers are each provided to a gate of a respective one of a plurality of pull-down transistors, and wherein the fourth status signal is generated by coupling drains of the pulldown transistors.

53. (New) The system of claim 4, wherein voltages for memory cells not selected for programming are set to approximately zero.

In the Drawings:

Figs. 4A-4C and 14 have been amended as indicated by the redlined revisions.

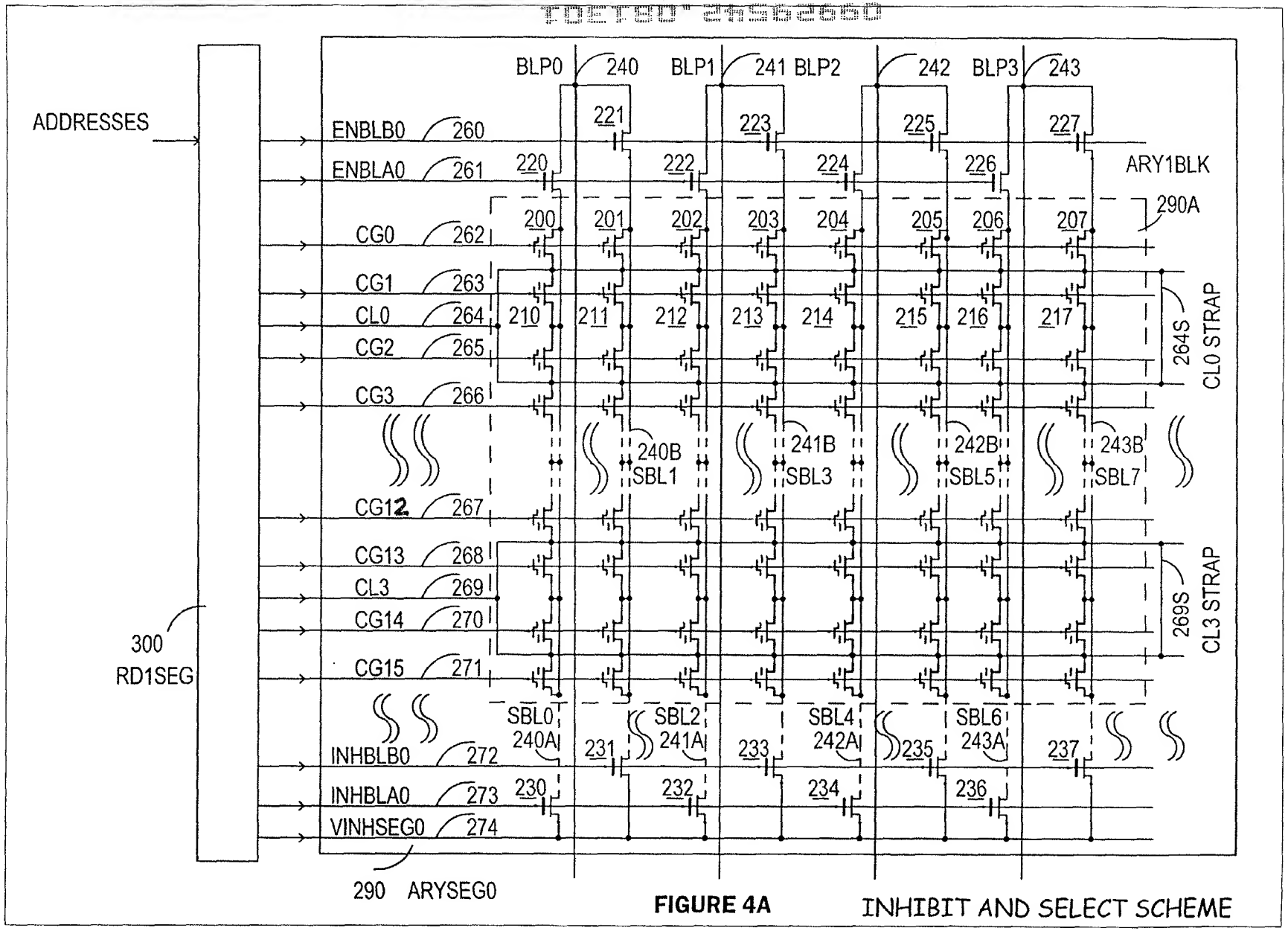


FIGURE 4A INHIBIT AND SELECT SCHEME

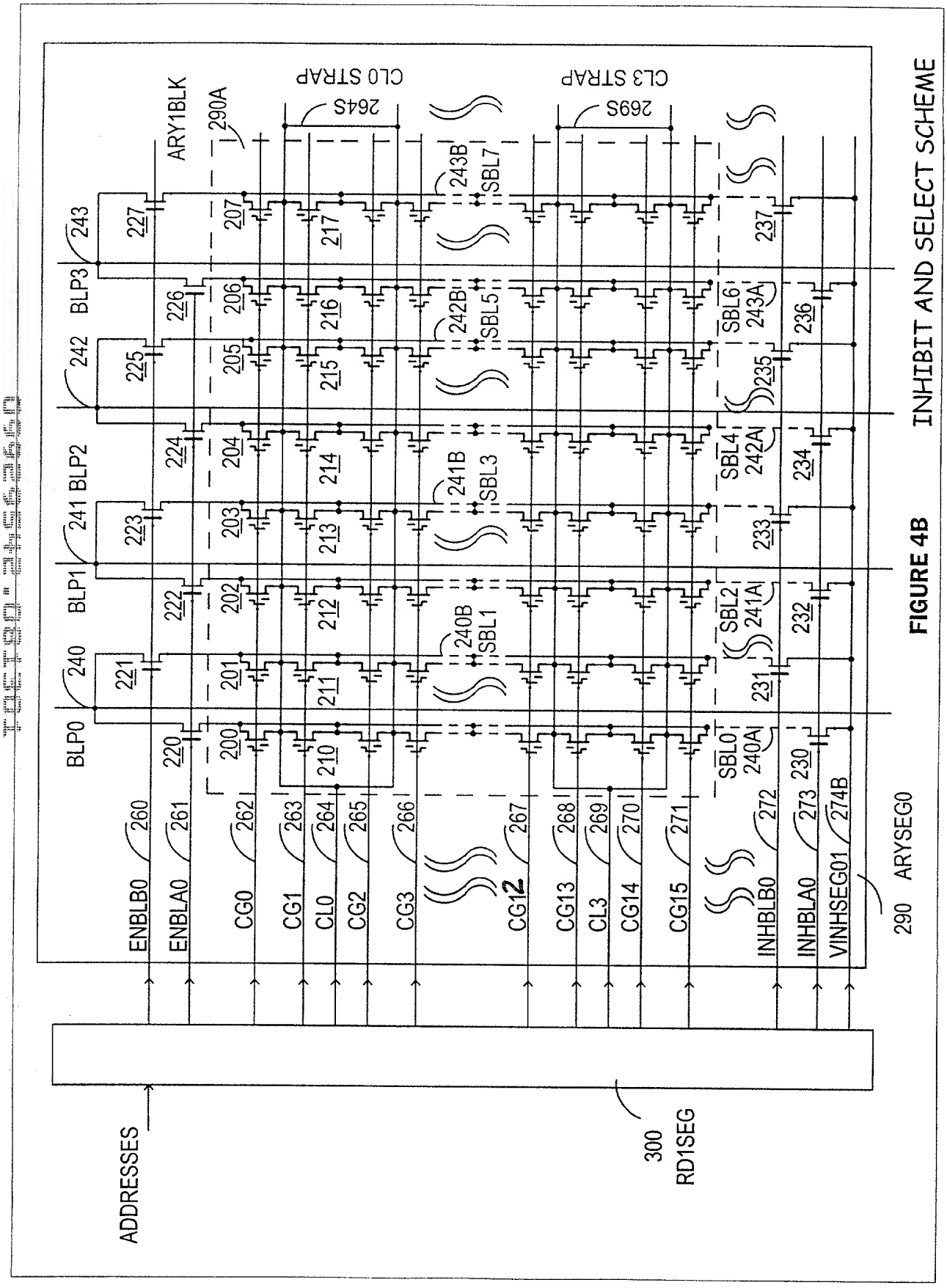


FIGURE 4B INHIBIT AND SELECT SCHEME

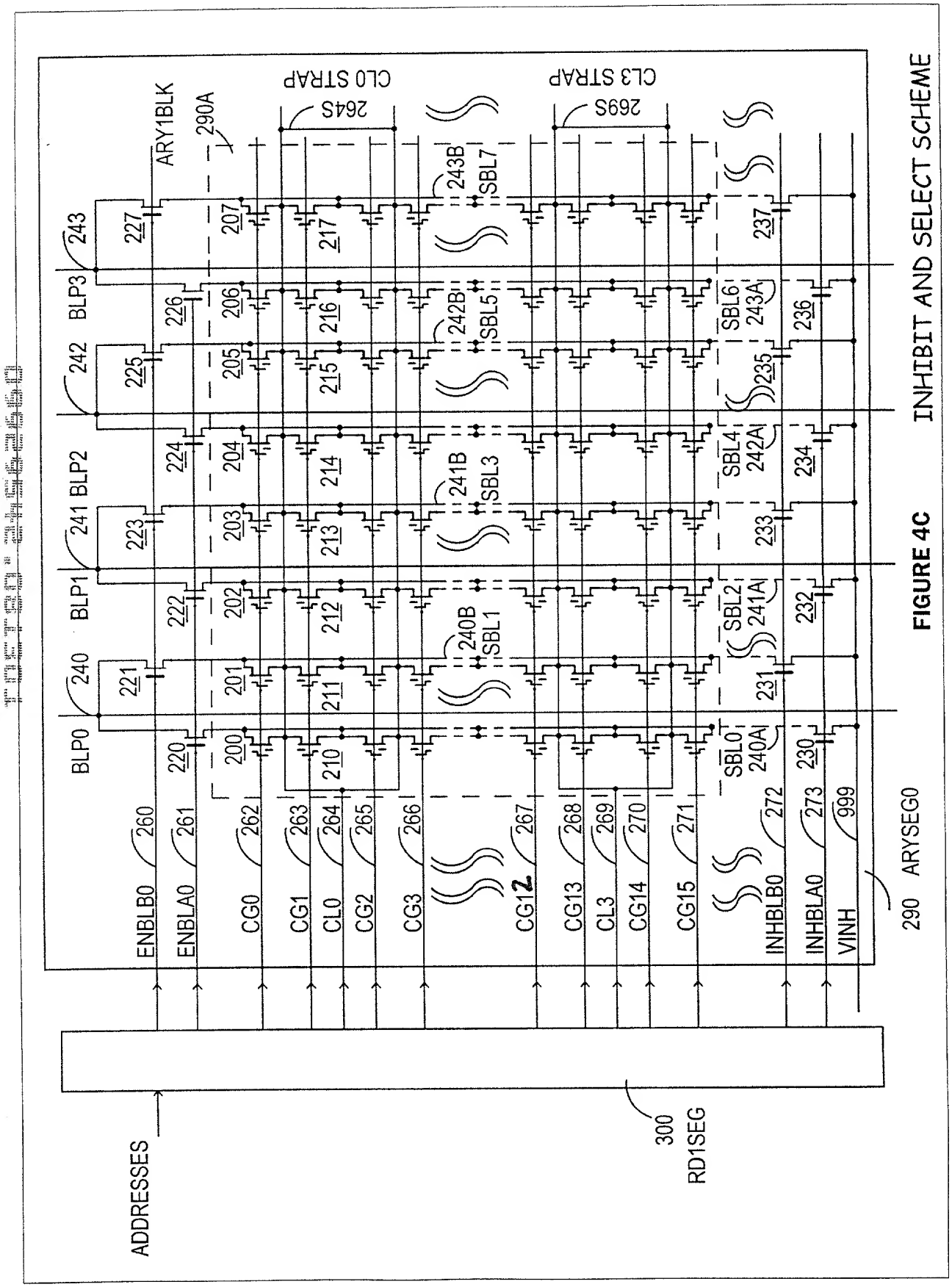
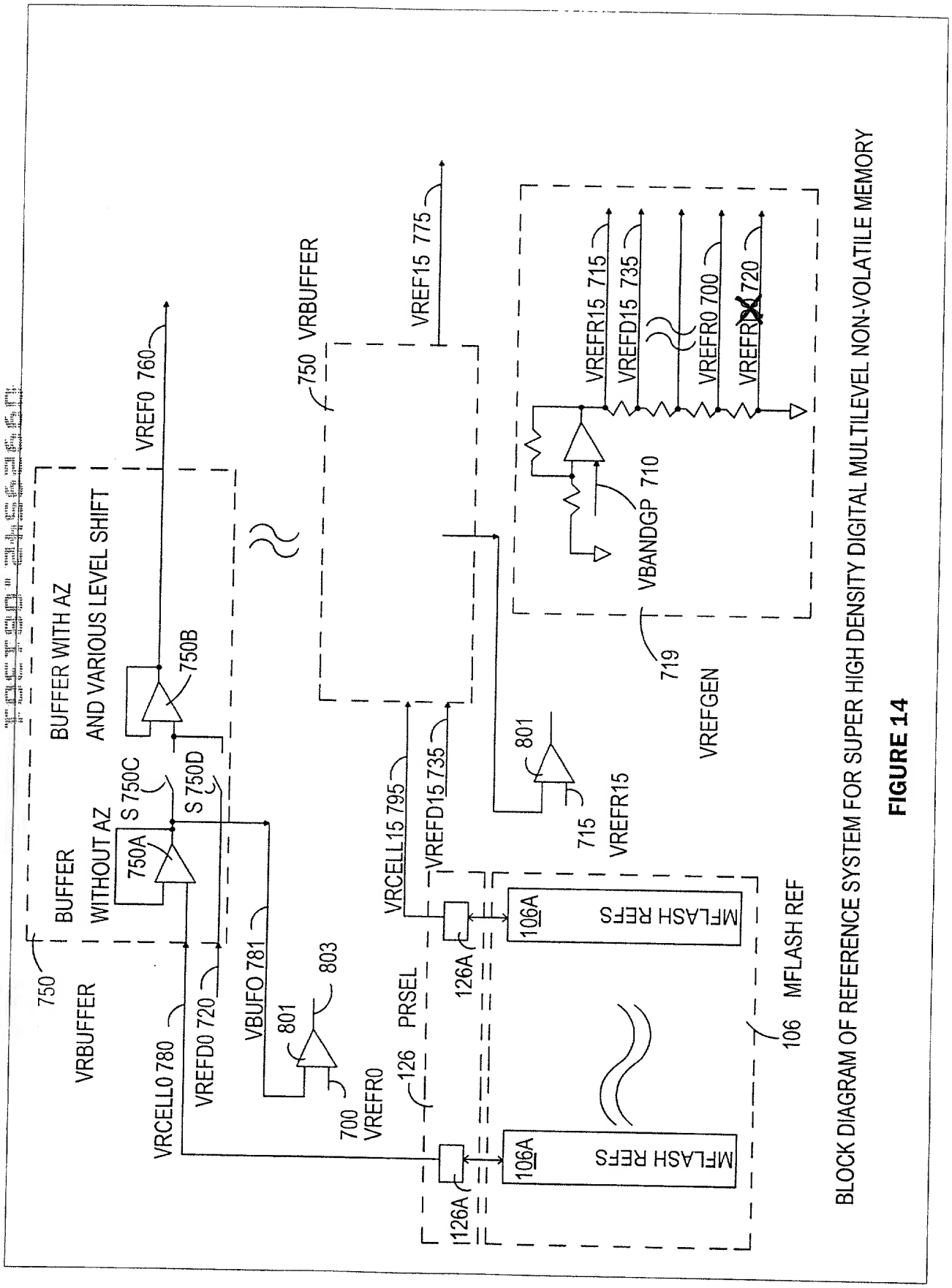


FIGURE 4C INHIBIT AND SELECT SCHEME



BLOCK DIAGRAM OF REFERENCE SYSTEM FOR SUPER HIGH DENSITY DIGITAL MULTILEVEL NON-VOLATILE MEMORY

FIGURE 14